ECONOMIC RESEARCH
FEDERAL RESERVE BANK OF ST. LOUIS

WORKING PAPER SERIES

The Beige Book and the Business Cycle: Using Beige Book Anecdotes to Construct Recession Probabilities

Authors	Charles S. Gascon, and Joseph Martorana
Working Paper Number	2024-037A
Creation Date	November 2024
Citable Link	https://doi.org/10.20955/wp.2024.037
Suggested Citation	Gascon, C.S., Martorana, J., 2024; The Beige Book and the Business Cycle: Using Beige Book Anecdotes to Construct Recession Probabilities, Federal Reserve Bank of St. Louis Working Paper 2024-037. URL https://doi.org/10.20955/wp.2024.037

Federal Reserve Bank of St. Louis, Research Division, P.O. Box 442, St. Louis, MO 63166

The Beige Book and the Business Cycle: Using Beige Book Anecdotes to Construct Recession Probabilities*

Charles Gascon Joseph Martorana

Abstract

The Federal Reserve releases the Beige Book prior to each Federal Open Market Committee meeting. The report is a narrative based on anecdotal and qualitative information collected from a wide range of contacts in each of the 12 Federal Reserve Districts. We take the lexicon approach to text analysis to create sentiment indexes that track changes in economic conditions from the very first Beige Book in May 1970 to the most recent (at the time of writing) in October 2024. We create additional indexes to account for various current-event shocks, such as political events or natural disasters that distort typical sentiment measures. We find that the real-time recession probabilities derived from a probit model featuring only the created sentiment and shock indexes are closely correlated with NBER recession periods, and more accurately indicate business cycle turning points than other widely cited measures. We find that the Beige Book can be used to promptly identify periods of economic recession as our model typically allows us to date business cycle turning points far in advance of the official announcements made by the National Bureau of Economic Research's Business Cycle Dating Committee.

Keywords: Beige Book, Federal Reserve System, Sentiment Analysis, Recession, Business Cycle.

JEL Codes: E3, E58

---

* Gascon is a Economist and Research Officer of the Federal Reserve Bank of St. Louis, and Martorana is a Research Assistant at the Federal Reserve Bank of St. Louis.
The authors thank attendees of the Federal Reserve System Committee on Regional Analysis for helpful comments.
The views expressed herein are those of the authors and do not reflect the ones of the Federal Reserve Bank of St. Louis, the Federal Reserve System, or the Board of Governors.

The Federal Reserve's Beige Book, formally known as the Summary of Commentary on Current Economic Conditions, is published eight times a year, typically two weeks prior to a Federal Open Market Committee (FOMC) meeting. It is a public report that summarizes regional changes in economic activity since the previous FOMC meeting. Unlike formal statistical reports, such as the monthly Employment Situation or the Federal Reserve's Industrial Production and Capacity Utilization reports, the Beige Book is qualitative and based on comments from business and community contacts. Representatives from each Federal Reserve Bank ask contacts, directors, or members of various "economic councils" within its District about topics like current and expected trends in employment, prices, and economic activity. The twelve District reports are presented alongside the National Summary, a summary of the balance of all District reports.[1]

The financial press provides fairly comprehensive reporting on each Beige Book release given its nature and its timing. Similarly, many economic analysts and financial market participants regularly dissect the report to gauge the economy's prospects and ultimately what they might imply for the future stance of monetary policy. However, some economists believe that the Beige Book's "soft" content is inferior to hard data.[2] Regardless, the report is intended to provide a timely assessment of national and regional economic conditions to complement other sources of information.

Over time, economists have adopted statistical methods of digesting qualitative information to quantify economic signals embedded within narratives, similar to the Beige Book. One of the earliest and best-known efforts to quantify the information contained in newspaper articles was the development of economic policy uncertainty indexes by Bloom (2009) and subsequently enhanced by Baker, Bloom, and Davis (2016). More recently, Handlan (2022A, 2022B) and Gati and Handlan (2023) use text analysis to measure how the wording within FOMC communications influence asset prices, how FOMC communications shift market expectations of future monetary policy, and how communication rules employed by the FOMC have evolved overtime with the economy.

This and other recent work have taken advantage of advanced software or neural networks, but there is also a broad literature in which text is analyzed more simply. While basic methods of text analysis may lack the nuance of large language models (LLMs), for example, such less-elaborate techniques have the benefit of being more accessible and easy to understand. In the lexicon approach to text analysis, an analyst counts the number of words in each Beige Book that convey positive or negative sentiment according to a predetermined "dictionary" or lexicon. These numerical counts are a useful addition to the toolkit economists use to track economic conditions. Specifically, this approach can be used in real time with the Beige Book to track

[1] District reports generally contain a District summary and several sections describing trends in different sectors. The National Summary generally contains a summary of national activity and another summary for each District. The layouts of the National Summary and District reports have changed several times since the Beige Book's inception in 1970, but are broadly structured as described.

[2] Former Federal Reserve Vice Chair Alan Blinder (1997) wrote "there are two basic ways to obtain quantitative information about the economy: you can study econometric evidence, or you can ask your uncle … I believe there is far too much uncle-asking in government circles in general and in central banking circles in particular."

business cycle turning points with impressive accuracy. Moreover, the lexicon approach allows observers to detect turning points far in advance of the official announcements made by the National Bureau of Economic Research's Business Cycle Dating Committee (NBER).

1. Beige Book Overview

Monetary policy makers have relied on anecdotal information to appraise economic conditions for decades. In 1970, then-FOMC chairman Arthur Burns created the monthly "Current Economic Conditions by District" to compile information on economic conditions across the Federal Reserve Districts prior to each FOMC meeting. Known informally as the Red Book for its red cover, the report was first published on May 20, 1970, for the May 26, 1970, FOMC meeting. As the introductory paragraph of the inaugural Red Book states:

> This initial report of economic conditions in the 12 Federal Reserve Districts is based on information gathered from directors of the Reserve Banks, conversations with local bankers, businessmen and economists, regular monthly surveys of manufacturing and trade industries conducted by some of the Reserve Banks, and selected statistical measures of regional economic activity.[3]

Throughout his tenure as chairman of the Subcommittee on Domestic Monetary Policy in the early 1980s, D.C. Congressman Walter Fauntroy pushed for more transparency in the FOMC's decision-making process. To this end, he requested that the documents prepared prior to FOMC meetings be made public. As a result, the Red Book was then made available to the public in 1983 following two significant changes. First, it was released two weeks prior to each FOMC meeting rather than one week prior. Second, the color of its cover was changed from red to beige and so became known as the Beige Book.[4] Further mention of "Beige Book" will refer to the history of publications since 1970, which includes Red Book publications prior to 1983.

Economists have increasingly attempted to assess the Beige Book's economic significance quantitatively through a variety of statistical techniques. Rolnick, Runkle, and Fettig (1999) conduct a study in which they read and score the National Summary of 265 Beige Books. They find that the report's "eyeball evidence" could be used to predict the growth of current-quarter real gross domestic product (GDP) with statistical significance but it does not improve upon private sector forecasts for national output. In their view, the Beige Book is "a better mirror than crystal ball."

Balke and Peterson (2002) conduct a similar study in which they read and score the Beige Book National Summary and District reports. They additionally create sector-level scores that capture developments in retail, manufacturing, banking, construction, and natural resource industries. They find that the average District report scores, the sector scores, and National

[3] See https://www.minneapolisfed.org/beige-book-reports/1970/1970-05-su

[4] For more information on historical materials of the Federal Reserve System: https://www.federalreserve.gov/monetarypolicy/fomc_historical.htm

Summary scores individually and jointly contain significant explanatory power for both current and next quarter real GDP growth. Further, they find that Beige Book indexes contain information that is not available in the Blue Chip Consensus Forecast and next quarter real GDP forecasts derived from a real-time time-series model using lagged real GDP, industrial production, employment growth, and real retail sales.

In an attempt to remove the subjectivity of reader-assigned scores, Armesto et al. (2009) use dictionary-based content analysis software to measure the optimistic and pessimistic tones of each Beige Book from May 1970 through July 2005, and a mixed data sampling model to account variation in the Beige Book's report frequency for predicting real output growth. They find that the usefulness of the Beige Book in predicting regional employment varies by District, with the variation of employment explained by each District's report ranging from 2% to 18%. Similarly, Balke, Fulmer, and Zhang (2017), create a quantitative Beige Book index using text analysis and incorporate the index into a factor model with six other economic indicators. They find the Beige Book scores provide additional information not contained in other real-time data for about three weeks after the report's publication.

Most recently, Filippou et al. (2024) create national and District sentiment indexes from the Beige Book and find that the sentiment indexes can be used to forecast and nowcast the probability of recessions. To read and quantify the Beige Books, they use FinBERT, a domain-adapted variation of Google's LLM, BERT, trained specifically on a corpus of financial text (Huang, Wang, and Yang 2023).[5] FinBERT and other LLMs are able to understand semantic and syntactic relations among words and are typically trained on large volumes of text. These cutting-edge models are able to read with context, meaning that a single part of a given text is not looked at in isolation but that information is extracted from the surrounding text and used in its classification or summarization. This is more like how a human reader would go through a text compared to less-advanced algorithms, and while this technique creates a richness within the Beige Book sentiment indexes, it comes at a cost of complexity. Large language models typically have millions of parameters and BERT is no exception. The base version of BERT, called $\text{BERT}_{\text{BASE}}$, contains 110 million parameters. A larger version of BERT, $\text{BERT}_{\text{LARGE}}$, contains 325 million parameters. This makes it difficult for an analyst to fully follow how such a model arrives at its conclusions given an input text.

An alternative approach is to consider a list of words that will always be positive or will always be negative, and to never consider context. Then, essentially count the positive and negative words that appear in the text to determine the sentiment expressed within the text. This lexicon approach, also known as the "bag of words" approach, though more rigid, is still relevant given that its rules are very straightforward and that its results are easy to replicate, even by hand. This simplicity lends the results to direct interpretation and understanding, as given any text and its score, anyone could easily determine why the text was deemed positive or negative

[5] BERT stands for Bidirectional Encoder Representations for Transformers, and was developed by Devlin, Chang, Lee, and Toutanova (2019). A thorough discussion of the technology underpinning BERT is outside the scope of this paper. Rogers, Kovaleva, and Rumshishky (2020), though, provide an overview of BERT and its contributions to natural language processing.

by the model. What's more, considering that results from both the lexicon and the LLM approaches are very similar, it does not seem that one methodology is obviously superior to the other in this application.[6] Lastly, while these methodologies represent two poles of sophistication, their concurring results suggests that, regardless of the tool used, the Beige Book does contain a fair assessment of economic conditions.

2. Data Summary and Methodology

2.1 Sentiment Indexing

All Beige Books, from the very first in May 1970 to the most recent of October 2024, are used to form a dataset that contains the text and date of each report.[7] After cleaning the Beige Books, their text is filtered through different dictionaries containing words that are labeled to indicate whether they convey positive or negative sentiment.[8] Six dictionaries in total are used, with three considered to be general, domain-agnostic dictionaries (Stone et al. 1966; Wilson, Wiebe, and Hoffman 2005), and the other three considered to be economic or finance-specific dictionaries (Henry 2008; Loughran and McDonald 2011; Barbaglia et al. 2023).

The distinction between general and economic dictionaries is important. As Loughran and McDonald show, words that typically carry negative connotations in general language often are not so negative in a financial context. They find that nearly 74% of the volume of words classified as negative according to the Harvard Psychosociological Dictionary (H4) in a sample of over 50,000 10-K reports do not have a negative financial meaning. They emphasize words like "cost," "tax," "foreign," "crude," "expense," and "liability" among many others as words that send different signals to different audiences, and propose a finance-adapted version of the H4, commonly referred to as the Loughran-McDonald Dictionary (LMD). Both the LMD and the H4 are used to measure Beige Book sentiment here.

Sentiment for a report at time t is then calculated according to each dictionary d using the equation

$$Sentiment_{d,t} = \left(\frac{P_{d,t} - N_{d,t}}{T_t}\right) * 100. \qquad \textbf{(1)}$$

[6] This is not to say that LLMs never possess significant advantages over the simpler algorithm we employ here. In many applications they do, and the lexicon approach would clearly not be suitable for tasks like summarization and location of certain topics within texts.
[7] Beige Book text is sourced from the Beige Book Archive housed on the Federal Reserve Bank of Minneapolis webpage.
[8] The reports are cleaned by removing, where possible, null words—words that are neither positive nor negative—in the text as these words contain no relative information about sentiment. Such words are typically articles, prepositions, pronouns, and auxiliary verbs. Additional null words that appear regularly in the report, such as "district," the names of the Federal Reserve Banks, the names of U.S states and U.S. geographic regions, and month names, are also mostly removed. Numbers are removed from the text as well. The remaining text is lemmatized to reduce each word to its root. For example, "decreasing," "decreased," and "decreases" are all distinct words but are reduced to "decrease" after lemmatization. This simplifies word matching in the text as distinct words that may represent the same idea and same sentiment are grouped together for analysis.

The total number of positive and negative words in the Beige Book contained in each dictionary are represented by P and N, respectively, while T represents the number of all words in the report. The three general sentiment scores and the three economic sentiment scores for each Beige Book are then averaged to create two aggregate sentiment indexes. Using multiple dictionaries in this way smooths some of the idiosyncratic movements in sentiment according to the individual dictionaries and reflects the changes in sentiment that are captured across dictionaries. The aggregate indexes, called the Average General Sentiment Index (AGS) and Average Economic Sentiment Index (AES), are created according to **Equations 2a** and **2b**

$$Average\ General\ Sentiment_t = \frac{Sentiment_{G1,t}+Sentiment_{G2,t}+Sentiment_{G3,t}}{3}, \text{ and} \qquad \textbf{(2a)}$$

$$Average\ Economic\ Sentiment_t = \frac{Sentiment_{E1,t}+Sentiment_{E2,t}+Sentiment_{E3,t}}{3}, \qquad \textbf{(2b)}$$

where G1, G2, and G3 represent the three general dictionaries and E1, E2, and E3 represent the three economic dictionaries.[9]

Figure 1 presents the AGS and AES (standardized with means of zero and unit standard deviations) overlayed with recession periods determined by the NBER. Ostensibly, the two indexes look very similar. They both exhibit substantial noise, but the largest deteriorations in sentiment typically accompany period of recession. Notably, at the end of recession periods the indexes quickly reverse their declines, even though their values may still remain below average levels.

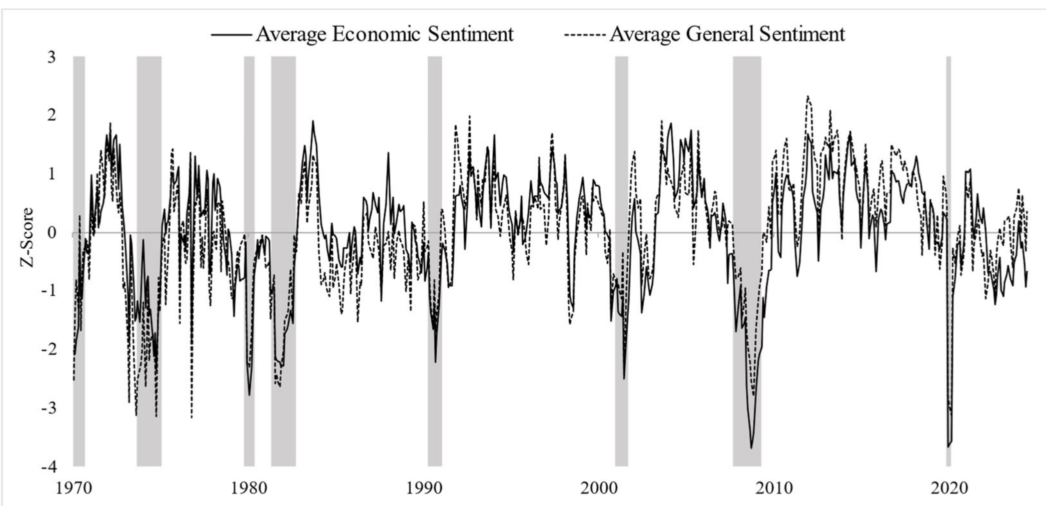


Fig 1. Historical Sentiment Scores

Notes: The AGS (solid) and AES (dashed) are standardized using their own averages and standard deviations over the entire sample period. Accordingly, 0 does not represent "neutral" sentiment but average sentiment expressed within the Beige Book over the several decades. Shaded grey bars represent NBER recession periods.

[9] A discussion the individual sentiment measures according to each dictionary is available in the Appendix. Differences in performance of each economic sentiment measure are immaterial, though the differences are more pronounced among the general sentiment measures.

However, a more careful look reveals subtle but material differences. First, the AES is more strongly above average during the late 1970s and the 1980s than is the AGS. Second, while the AGS shows that sentiment fell to roughly 3 standard deviations below average during the 1973 recession, global financial crisis (2008) and the pandemic recession (2020), the AES shows that sentiment fell only to 2 standard deviations below average in 1973, but 4 standard deviations below average in 2008 and 2020. The indexes have also diverged in recent years. As of the latest Beige Book, there is roughly a 1 standard deviation spread between the AGS (0.36 above average) and the AES (0.67 below average). That the AES shows sentiment has been below average for most of the last two years is more in line with other popular sentiment measures, like the University of Michigan's Consumer Sentiment Index, which has shown consumer sentiment to be much lower in 2024 than during the depths of the 2020 recession (Gascon and Martorana, 2024).

Given that the spread between the AGS and AES scores can be quite substantial, a few points are worth highlighting from a comparison of the scores according to the LMD and the H4 for the October 2024 Beige Book. The discussion of retail activity in the Dallas report is particularly polarizing. That section has the fourth lowest score of all 126 subsections in the current Beige Book according to the LMD, but is in the top third of highest scores according to the H4. The section is provided below according to the H4 on the left and the LMD on the right.[10] Negative words are underlined with a wave while positive words are double-underlined.

According to H4:

Retail sales **fell** during the past six weeks. Auto dealers reported slow traffic and **declining** sales. However, not all sectors **experienced** a slowing. **Health** and personal **care** retailers cited **modest** increases, and food and beverage stores generally reported **flat** activity. Retail inventories held **steady**. Outlooks remained **negative**.

According to LMD:

Retail sales fell during the past six weeks. Auto dealers reported slow traffic and **declining** sales. However, not all sectors experienced a slowing. Health and personal care retailers cited modest increases, and food and beverage stores generally reported **flat** activity. Retail inventories held **steady**. Outlooks remained **negative**.

The H4 recognizes a few words as positive though they do not reflect any sentiment at all, such as "experienced," "health," and "care." It is easy to see how in a typical day-to-day context, these words would be seen as positive, but in this context they skew sentiment upward. The H4 also notes that "modest," "flat," and "steady" are positive while the LMD considers these words to be negative. The LMD does not mark any words in the text as positive, and marks "slow" and "slowing" as additional negative words. The end result according to the two dictionaries is a section that is either overwhelmingly positive or exclusively negative.

[10] See https://www.minneapolisfed.org/beige-book-reports/2024/2024-10-da

This does not mean that economic dictionaries always provide scores that are perfect representations of the sentiment expressed within the text. Consider the following few sentences of the labor markets subsection of the latest Atlanta report, annotated according to the LMD:

> Firms reported plans to keep headcount roughly flat for the remainder of the year, and while a growing minority reported modest and targeted reductions in headcount in response to **slowing** demand, most were not planning to implement widescale **layoffs**. Some firms reported welcoming or **encouraging** **attrition**, hesitating to backfill positions, or reducing hours. A minority of firms were hiring for growth.[11]

The LMD identifies more negative than positive words in the given block of text, and rightfully so. Typically, "attrition" and "layoffs" would be considered negative. However, in this case contacts have suggested that they are avoiding layoffs and view attrition favorably. The H4 sentiment (not shown here), is further from the mark because it detects no negative words in these sentences and treats "flat," "modest," "welcoming," and "encouraging" as positive. Ultimately, a human reader would likely disagree with the rating provided by the H4; the same reader would likely be in agreement with the LMD's conclusion, but not its reasoning.

2.2 Shock Indexing

Because the Beige Book synthesizes regional anecdotes collected from contacts in each Federal Reserve District, it contains references to extraneous events that may be distinct from broad, underlying economic trends but influence sentiment nonetheless. Accordingly there are some instances where the sentiment indexes are pushed downward but are unaccompanied by recessions. To correct for this, two indexes are created according to two lists of key terms to track the shocks reflected in each report. The Physical Index reflects the frequency of occurrence of words like "hurricane," "earthquake," "oil spill," "chemical spill," and others in order to isolate mentions of environmental or natural disasters within the text. The Political Index similarly represents the wars, campaigns, riots, labor strikes, and terrorist attacks that anecdotes have referenced throughout the Beige Book's history. The word lists used for each shock index are provided in the Appendix.

The shock indexes are constructed using the formula

$$Shock_{L,t} = \left(\frac{K_{L,t}}{T_t}\right) * 100, \tag{3}$$

where K is the number of keywords from word list L contained in each Beige Book at time t, and T is the total number of words in the report. The result is the percentage of words in each report that fall in the lists of exogenous shock keywords ($Shock_{L,t}$). It is worth emphasizing that these shock indexes are not sentiment indexes, but their inclusion in the model helps (in theory) to

[11] See https://www.minneapolisfed.org/beige-book-reports/2024/2024-10-at

distinguish changes in sentiment related to current events from changes in sentiment related to economic trends.

The shock indexes, selectively annotated, are presented in **Figure 2**. They do reasonably well at identifying major events discussed in the reports. The Physical Index reaches its highest points around the times of Hurricanes Katrina, Sandy, and Harvey in 2005, 2012, and 2017, respectively. The Great Flood of 1993 also delivered a sizeable shock, and the 1988 North American Drought kept index levels slightly elevated through its end in 1990. The Political Index spikes immediately following the events of September 11th in 2001 and then again at the beginning of the Iraq War in 2003. Truckers' strikes in the 1970s also pushed the index to impressive heights.

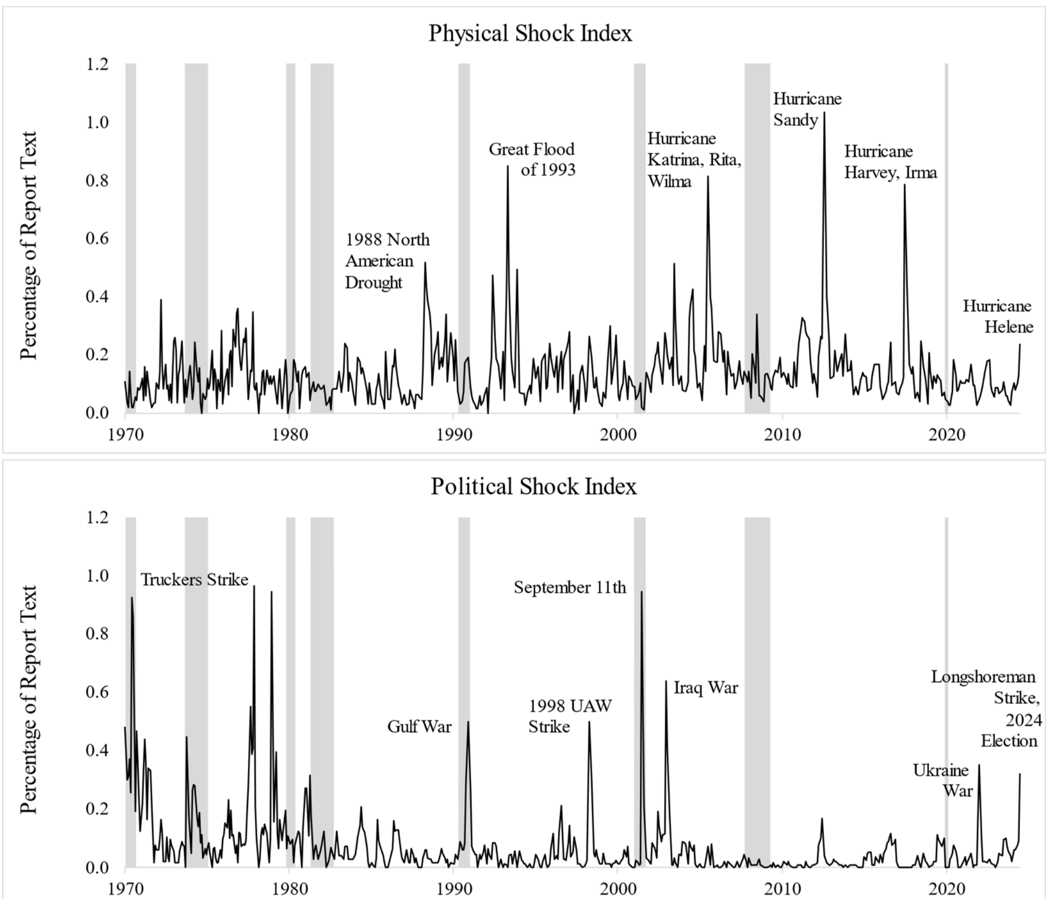


Fig 2. Historical Sentiment Scores

Notes: The shock indexes have been selectively annotated to highlight major events that the Beige Books have captured over the past decades, and this list of events is not exhaustive. Shaded grey bars represent NBER recession periods.

Simple least squares regressions of the sentiment indexes on the shock indexes show that the shocks exert considerable influence. The effects of each shock on the AGS and AES are both economically and statistically significant, as shown in **Table 1**, though the variations in the changes of each shock explain very little of the variation of the changes in overall sentiment. As the Physical Index increases by one percentage point, for example, the AGS falls by 1.05

standard deviations. The same increase in the Political Index has a similar impact on the AGS, causing it to fall by 0.95 standard deviations. When the shock indexes are considered jointly (column 3), their effects on the AGS are slightly greater. Their individual effects on the AES, though, are somewhat less. Still, in some cases the influence of the shocks can be clearly identified in **Figure 1**. The shock associated with September 11th, for example, translates to a 1.00 standard deviation reduction in the AGS and a 1.20 standard deviation reduction in the AES. Both of these declines are apparent immediately before the end of the dot-com recession. Hurricane Sandy also coincides with a sharp drop of 1.40 standard deviations in the AGS and of 0.8 standard deviations of the AES in October 2012.

Dependent variable:	(1) AGS	(2) AGS	(3) AGS	(4) AES	(5) AES	(6) AES
Physical Shock	-1.049*** (0.292)		-1.104*** (0.292)	-0.821*** (0.229)		-1.412*** (0.378)
Political Shock		-0.947** (0.344)	-0.993** (0.348)		-0.665** (0.257)	-1.153** (0.424)
Constant	0.013 (0.031)	0.012 (0.031)	0.015 (0.030)	0.002 (0.025)	0.001 (0.025)	0.006 (0.040)
Adjusted R^2	0.026	0.027	0.055	0.024	0.020	0.046
Observations	472	472	472	472	472	472
F-Statistic	12.95***	7.55**	10.69***	12.85***	6.71**	9.74***
(P > F)	(0.000)	(0.006)	(0.000)	(0.000)	(0.001)	(0.000)

Table 1: In-sample Effect of Shocks on Sentiment

Notes: Sentiment scores are standardized and first-differenced, while shock indexes are only first-differenced. Coefficients represent standard deviation changes in sentiment per percentage point change in shock indexes. Robust standard errors are in parentheses except where noted; *** p<.001, ** p<.01, * p<.05.

2.3 The Probit Model

Following Estrella and Mishkin (1998), Wright (2006), Chen, Iqbal, and Lai (2011), Hao and Ng (2011), and others, a probit model is used to map the sentiment and shock indexes to out-of-sample probabilities of current-period recession. The model is given by **Equations 3a** and **3b** where P_{AGS} and P_{AES} represent the probability of a contemporaneous recession according to the AGS and AES. The binary, real-time recession indicator at report date t is given by R_t and X is the collection of all available information at that time:

$$P_{AGS}(R_t = 1 \mid X_t) = \phi(\beta_0 + \beta_1 AGS_t + \beta_2 AGS_{t-1} + \beta_3 E_{Pol,t} + \beta_4 E_{Phy,t} + \varepsilon_t), \text{ and} \quad \textbf{(3a)}$$

$$P_{AES}(R_t = 1 \mid X_t) = \phi(\beta_0 + \beta_1 AES_t + \beta_2 AES_{t-1} + \beta_3 E_{Pol,t} + \beta_4 E_{Phy,t} + \varepsilon_t). \quad \textbf{(3b)}$$

The current and lagged measures of the AGS are given in **Equation 3a** by AGS_t and AGS_{t-1}, respectively, and similarly for the AES in **Equation 3b**. Variables $E_{Pol,t}$ and $E_{Phy,t}$ represent the shock indexes at time t as well. The standard normal error term of regression is given by ε_t. The cumulative standard normal distribution function ϕ(.) is applied to the linear model. The result is a predicted probability value of a real-time recession occurrence bounded between zero and one.

3. Results

3.1 In-Sample Effects of Sentiment and Shocks

Table 2 shows the average partial effects of the standardized explanatory variables using the full sample available (i.e. in-sample estimations). The current and lagged effects of each sentiment index are first estimated in isolation and then supplemented with the two shock indexes. The coefficients and statistics presented in columns 3 and 6 correspond with the models laid out by **Equations 3a** and **3b** respectively.

Three points are worth noting. First, the total effects of the AGS and AES are almost equivalent and broadly significant. As shown in column 3, increases by one standard deviation in the lagged and current AGS reduce the probability of recession by 6.5 to 7.5 percentage points. The spread between the lagged and current AES, in column 6, is larger. A standard deviation movement in the current AES moves the probability of recession by 9 percentage points in the opposite direction, while the same increase in the lagged AES lowers the probability by five percentage points on average. Regardless of which sentiment measure is used, an increase in Beige Book sentiment translates to a material and statistically significant decline in the probability of recession.

Dependent variable:	(1) P_{AGS}	(2) P_{AGS}	(3) P_{AGS}	(4) P_{AES}	(5) P_{AES}	(6) P_{AES}
Sentiment	-13.47*** (1.21)		-7.47*** (1.51)	-14.77*** (1.25)		-9.20*** (1.50)
Sentiment Lagged		-13.38*** (1.25)	-6.52*** (1.33)		-13.96*** (1.36)	-5.07*** (1.74)
Physical Shock			-3.99* (1.59)			-2.08 (1.22)
Political Shock			0.96 (0.68)			1.56* (0.62)
Observations	472	472	472	472	472	472
χ^2 (P > χ^2)	66.7 (0.000)	66.5 (0.000)	61.8 (0.000)	35.8 (0.000)	42.6 (0.000)	35.8 (0.000)
Pseudo R^2	0.430	0.405	0.520	0.552	0.471	0.598
Log Likelihood	-105.9	-109.3	-88.2	-83.2	-97.2	-73.9
Akaike Inf. Crit.	215.8	222.6	186.4	170.4	198.3	157.8

Table 2: In-sample Effects on Recession Probability

Notes: The sentiment and shock indexes have been standardized according to their historical (1970 – 2024) means and standard deviations. Average partial effects are presented in percentage points. Coefficients represent standard deviation changes in sentiment per percentage point change in shock indexes. Robust standard errors are in parentheses; *** p<.001, ** p<.01, * p<.05.

Second, the shock indexes are considerably less influential in the probit model. In both models, an increase in the Physical Index leads to a decrease in the probability of recession. However, physical shock effects are only marginally significant when accompanied by the AGS. This is likely because the largest physical events most often occur during economic expansions. The index only budged slightly in response to Hurricanes Ike and Gustav in 2008, and that represents the largest physical shock registered during a recession. The four biggest physical events, though, all came during periods of growth.

Conversely, an increase in the Political Index leads to an increase in the probability of recession. But in the same way the physical shocks are generally insignificant, the political shock effects are only marginally significant when accompanied by the AES. The political shock index has the greatest coincidence of increases and recessions, the most notable of which are the events of September 11th and the dot-com recession of 2001. Other mild spikes in this index include the 1991 Gulf War, occurring alongside the recession of 1990 – 1991, and the truckers' strike of 1974.

Lastly, overall fit of the model to recession periods using the AES is largely better than the model using the AGS. Economic sentiment and the shock indexes account for just below 60% of the variability in the recession probability, according to the adjusted R^2. General sentiment and shocks account for 52% in comparison. Additionally, economic sentiment maximizes the log likelihood and minimizes the Akaike Information Criteria, suggesting a better fit to the probabilities over general sentiment.

3.2 Out-Sample Recession Probabilities

The models' real-time predictive capabilities can be demonstrated well by a recursive ex-ante analysis. The Beige Books from the first in 1970 through January 1979 are used to establish a baseline of coefficients for **Equations 3a** and **3b**.[12] Then beginning with the Beige Book from January 1980, the sentiment and shock index scores are used to obtain the contemporaneous probability of recession with the estimated coefficients. For each month going forward, the sample period for coefficient estimation is expanded by one month. In other words, all available information through time *t-12* are used to establish coefficients for predicting the probability of recession at time *t* using the Beige Book at time *t*.[13] The observations between time *t-1* and *t-11* are excluded in estimating the coefficients to allow sufficient time for the NBER to announce the most recent business cycle turning point, which typically happens within one year of its occurrence.

[12] Because the Beige Book has not been published monthly since 1983, the majority of the reports cannot be mapped directly to regular monthly data releases and the information contained in the reports does not readily correspond to any one month. We follow the approach outlined by Balke and Peterson (2002), in which the reports are assigned to a month according to the printed date through which anecdotes have been collected for the report. If the marked date occurs before the 15th of the month, we consider the Beige Book to contain information primarily for the previous month; if the marked date occurs on or after the 15th, we consider the report to be for that current month. For all reports published before 1983 (i.e. the monthly reports), we assume that the text relates to the month prior to the month of publication.

[13] In cases where a Beige Book is not available for time *t* when estimating the coefficients, we use the time-weighted average of the sentiment values of the last available and next available Beige Books.

The smoothed real-time P_{AGS} and P_{AES} obtained according to **Equations 3a** and **3b** are plotted with NBER recession periods in **Figure 3**. While both sets of probabilities change quickly around business cycle turning points, the AES probabilities clearly outperform their AGS counterparts. This is not surprising considering its superior measures of fit presented in **Table 2**. Most notably, the AGS probability of recession in the 1980s remains elevated while the AES probability stays much lower. This aligns with the fact that the AGS never moves much above its historical average during this period, unlike the AES, as discussed in **Section 2.1**. Furthermore, the AGS probabilities of recession in 2000 and in 2008, increase slightly but do not reach appreciably higher levels until well past the midpoint of the recessions. Conversely, the AES probability quickly surpasses 50% within the two months of the business cycle peak (even though some of those increases are partially retraced shortly after).

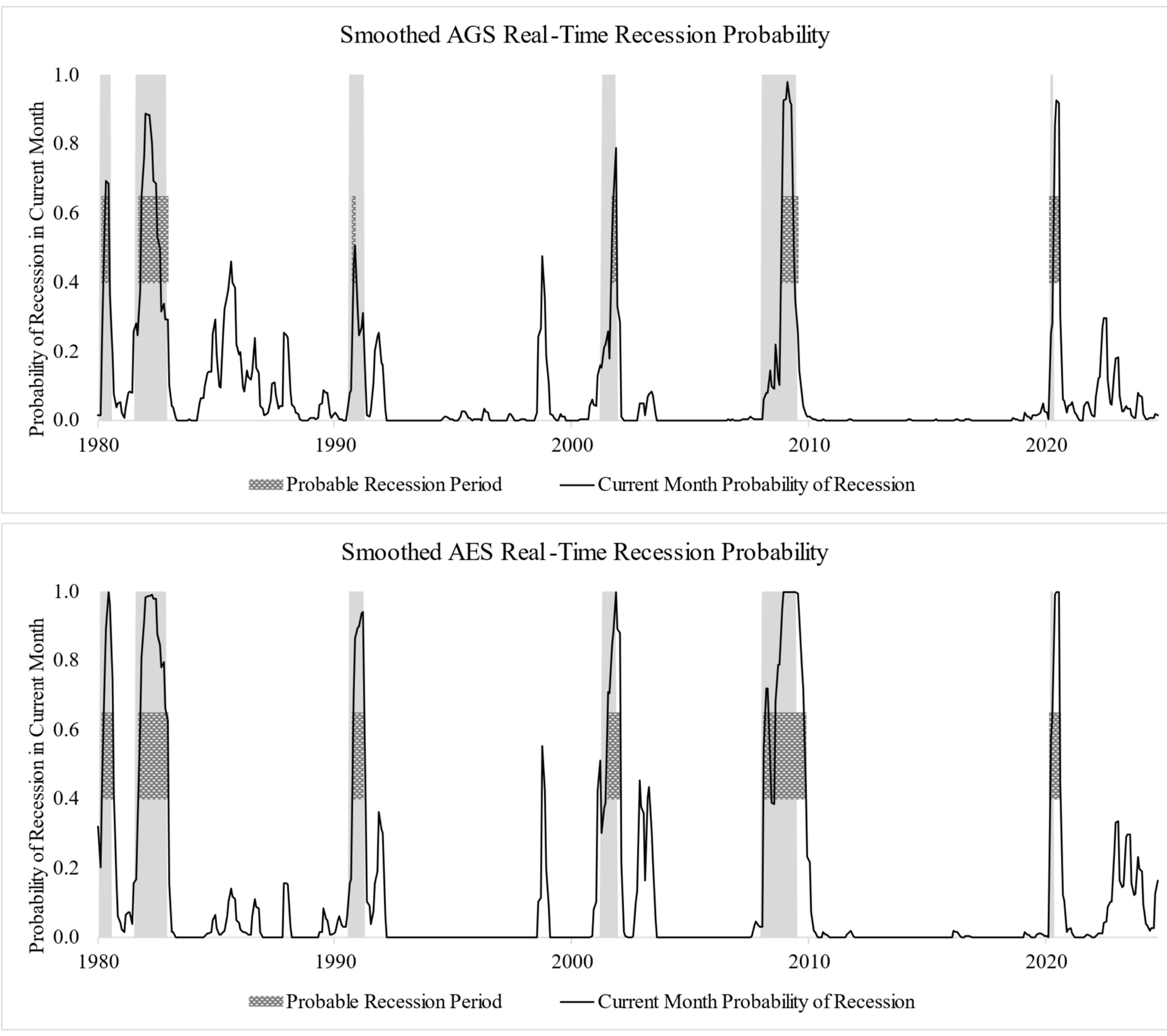


Fig 3. Real-Time Current-Month Probability of Recession

Notes: The time-weighted average probabilities have been presented for graphical purposes; the current observation receives a weighting of 1, the second observation receives a weighting of 0.67, while the third (earliest) observation receives a weighting of 0.33. Hashed-gray shaded regions represent the period over which the economy is likely in recession, according to the raw, unsmoothed probabilities and the rules presented in Equations 4a and 4b.

The unsmoothed P_{AGS} and P_{AES} from **Equations 3a** and **3b** can easily be transformed into periods in which a recession is likely to have occurred. To account for noise in the fitted probabilities, distinct rules are used to classify a period as a recession or an expansion. The rules are based on the work of Chauvet and Piger (2008) and make use of probability thresholds. The binary recession indicators according to the AGS and AES derived probabilities, R_{AGS} and R_{AES}, are given by the piecewise **Equations 4a and 4b** where a value of 1 indicates a recession in the current period and value of 0 indicates an expansion in the current period:

$$R_{AGS,t} = \begin{cases} 1 & if\ 0.40 \le P_{AGS,t}, P_{AGS,t-1} \le 1.00 \\ 0 & if\ 0.00 \le P_{AGS,t} \le 0.20 \\ R_{AGS,t-1} & if\ 0.20 < P_{AGS,t} < 0.40 \end{cases}, \text{ and} \tag{4a}$$

$$R_{AES,t} = \begin{cases} 1 & if\ 0.50 \le P_{AES,t}, P_{AES,t-1} \le 1.00 \\ 0 & if\ 0.00 \le P_{AES,t} \le 0.25 \\ R_{AES,t-1} & if\ 0.25 < P_{AES,t} < 0.50 \end{cases}. \tag{4b}$$

In simple terms, **Equation 4b** says that a recession occurs when P_{AES} is greater than or equal to 50% for two consecutive Beige Books. The recession is considered to have begun in the earlier of those months. An expansion occurs so long as the probability of recession is less than or equal to 25%. For all cases where P_{AES} is either greater than 25% but less than 50%, the current period is considered unchanged from the previous period.[14] If, for example, four Beige Books with P_{AES} above 50% are followed by a fifth report with P_{AES} of 30%, the latest report is still deemed recessionary.[15] The interpretation of **Equation 4a** is different only in that the thresholds for P_{AGS} are lower to reflect that the probabilities provided by **Equation 3a** are generally lower than those provided by **Equation 3b**.

The probable recession periods according to P_{AGS} and P_{AES} are shaded in hashed-gray in their respective panels of **Figure 3**. There is little overlap between the official NBER recessions and those identified by P_{AGS}. The P_{AGS} model misses almost all of the 1990 and dot-com recessions, and about half of the global financial crisis. By comparison, the turning points taken from the P_{AES} model are typically occur within one or two months of the observed turning points. The major exception is the global financial crisis where the model does not suggest the recession ended until November 2009, five months after the NBER end date.

It should be noted that while the P_{AES} model does not generate any false recessions, it comes close on two occasions. But even such near misses are not completely erroneous. In October 1998, the probability of recession spikes to 68% but then falls the following month to 25%, so no

[14] This rule is only a rule of thumb. A slightly different rule in which the thresholds are changed from 50% and 25% to 47% and 26%, respectively, would increase the number of correctly identified periods (recession and expansions) by 3.

[15] Months for which Beige Books are not available take the status of the most recent Beige Book.

recession is called based on **Equation 4b**.[16] Negative sentiment in the October Beige Book was primarily lead by the energy, agricultural, and manufacturing sectors. Contacts in the Dallas report noted that oil drilling activity was near "all-time low levels, with natural-gas directed drilling, offshore activity and international activity all declining slowly." Agricultural contacts in Minneapolis, Kansas City, Dallas, and San Francisco reported weakening profitability, low selling prices, and a "generally bleak pattern in agriculture." Lastly, manufacturers in Boston, Richmond, Dallas, and San Francisco reported a combination of weakening demand, lower orders, falling profitability, and layoffs, temporary furloughs, or plant closures (though some contacts in these Districts did confirm favorable conditions). The anecdotes relating to energy and agricultural are not so readily corroborated by available data, though the consumer price index for energy in U.S. cities had generally been declining over the previous year. Comments related to manufacturing, though, do align with the 1% decline registered in nominal manufacturers' new orders over that October and the generally flat level of activity that had endured over the previous year.

Similarly, in November 2002 and March 2003, P_{AES} reaches 54% and 59% but no recession is called given the interim period of lower probabilities.[17] Slow growth in retail sales, especially of automobiles, a soft labor market, and general uncertainty associated with the onset of the Iraq War weighed heavily on sentiment expressed in those Beige Books. Roughly half of the Districts in both the November 2002 and March 2003 Beige Books mentioned recent or imminent layoffs, or weakening regional demand as a result of layoffs. The National Summary of the November 2002 Beige Book further reported that "labor markets continued to be soft in nearly all Federal Reserve Districts." The anecdotes aligned with the continued decline of nonfarm payrolls long past the end of the dot-com recession and well into the second quarter of 2003. These weak economic conditions understandably coincided with relatively elevated probabilities of recession.

3.3 Alternative Comparisons

The stark difference in performance between P_{AGS} and P_{AES} suggests that dictionaries adapted to economic or financial texts do indeed carry significant advantages over broad, generic dictionaries. For this reason, this section focuses exclusively on the more successful P_{AES} model, whose ability to detect business cycle peaks is remarkably comparable to those of alternative methods and surveys. **Table 3** presents the accuracy and timeliness of the P_{AES} model alongside that of the real-time Sahm rule (Sahm 2019).[18] The first month of the last six NBER recessions are shown with their announcement dates, with the first month and date of identification shown according to the Sahm rule and the P_{AES}. In all cases, the Beige Book sentiment model identifies the recession before the official NBER turning point is announced. Additionally, the model's

[16] The Beige Book published in November 1998 corresponds with the October 1998 probabilities. See https://www.minneapolisfed.org/beige-book-reports/1998/1998-11-su

[17] The Beige Books published in November 2002 and April 2003 are aligned with the November 2002 and March 2003 probabilities. See https://www.minneapolisfed.org/beige-book-reports/2002/2002-11-su for the November 2002 Beige Book.

[18] The Sahm rule suggests a recession is likely underway when the three-month average national unemployment rate rises by at least 0.50 percentage points relative to its low over the previous 12 months.

estimated turning point and the turning point per the Sahm rule are usually within one month of each other.

The P_{AES} model, though, is less timely than the Sahm rule because of the irregular pattern of Beige Book releases. Take the dot-com recession as an example. The NBER announced in November 2001 that the recession began in April 2001. The Sahm rule would have been triggered that July with the release of the June employment report, predating the NBER's announcement by 143 days. By comparison, the Beige Book would have suggested that the recession began in July 2001, but this would not have been known until its release in September 2001. As a result, the recession would have been signaled 75 days after Sahm rule had already done so, even though the suggested start date of the recession differs from that of the Sahm rule by only 30 days.[19]

NBER Recession Start	NBER Recession Declared	Sahm Recession Start	Discrepancy (Months)	Sahm Recession Declared	Lead (Days)	P_{AES} Recession Start	Discrepancy (Months)	P_{AES} Recession Declared	Lead (Days)
Feb-80	6/3/1980	Apr-80	2	5/2/1980	32	Mar-80	1	5/14/1980	20
Aug-81	1/6/1982	Nov-81	3	12/4/1981	33	Oct-81	3	12/16/1981	21
Aug-90	4/25/1991	Nov-90	3	12/7/1990	139	Oct-90	2	12/5/1990	141
Apr-01	11/26/2001	Jun-01	2	7/6/2001	143	Jul-01	3	9/19/2001	68
Jan-08	12/1/2008	Apr-08	3	5/2/2008	213	Feb-08	1	4/16/2008	229
Mar-20	6/8/2020	Apr-20	1	5/8/2020	31	Mar-20	0	5/27/2020	12

Table 3: Comparison of Monthly Real-Time Recession Indicators

Notes: Discrepancy shows the number of months between the first actual and first perceived month of recession according to the real-time Sahm rule or P_{AES} model. Lead shows the days between the time at which a recession is first declared according to the Sahm rule or P_{AES} model (the Beige Book's publication date), and the date when the business cycle turning point is officially announced by the NBER. The declared dates according to the Sahm rule are the dates on which the employment situation report was released and the unemployment rate rose the 0.50 percentage points specified by the Sahm rule.

A similar analysis of the model's quarterly recession predictions in comparison with those of the Survey of Professional Forecasters (SPF) follows below, with the probabilities from each plotted in **Figure 4**. Both series display nearly identical peaks in the probability of recession, including the peaks that occur during expansions, although the SPF is considerably noisier than the P_{AES}. Regardless, both measures rise and fall just as quickly around business cycle turning points. The SPF and P_{AES} have tracked each other especially closely over the recent few years from 2021 through 2024. The probability of a recession according to both measures peaked between 30% and 40%, and was 16% as of the third quarter of 2024.

[19] If Equation 4b is modified as mentioned in Footnote 14, then the recession would have been considered to have begun in May, rather than July. This would have been apparent with the publication of the Beige Book on August 8th, increasing the lead time by 42 days.

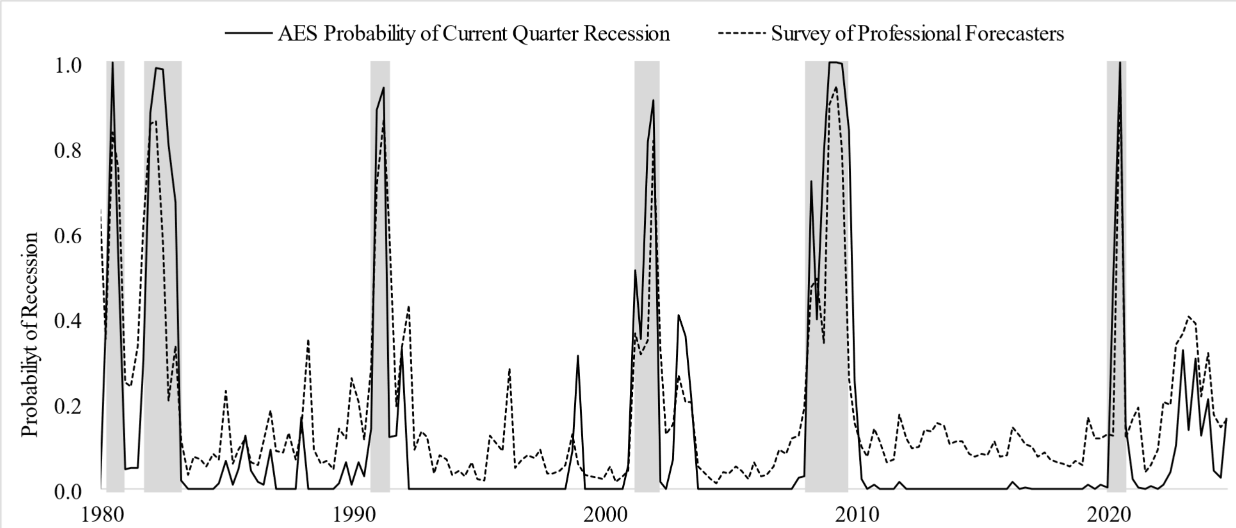


Fig 4. Real-Time Current Quarter Probability of Recession

Notes: The measure of probability taken from the Survey of Professional Forecasters (dashed) is respondents' assessment of the probability of negative real-GDP growth in the current quarter. AES Probability of Current Quarter Recession (solid) is calculated as the average of P_{AES} for each Beige Book in the given quarter.

Table 4 presents contraction and expansion starting dates for the SPF and P_{AES}.[20] The probability threshold for quarterly recession dating remains 50%, and the threshold for calling an expansion is again 25%. The thresholds for the SPF are slightly different, with the probability of a recession required to surpass 47% but stay below 26% for expansions.[21]

The table confirms what **Figure 4** illustrates. Both the SPF and the sentiment model identify the first quarter of a recession consistently well, and in all cases again the P_{AES} model identifies the recession in advance of the NBER announcements. The sentiment model's ability to detect expansions is just as impressive. It overestimates the length of only the global financial crisis by one quarter, but despite the late detection it still indicates that the recession had ended 250 days before the NBER announcement. While there are some instances in which either the SPF or P_{AES} model outperforms the other in terms of accuracy and timing, on balance they are roughly equivalent. Either can be used to supplement the other with reasonable confidence.

[20] The rule for quarterly recession dating according to the P_{AES} is slightly different from the rule used for the monthly recession dating. Because of the unusual pattern of Beige Book releases for most of its history, there is sometimes only one Beige Book released for a given quarter (though there have been two for each quarter for the last several years). As a result, it is not possible to seek out two consecutive reports in every quarter to determine whether a recession occurred. Additionally the monthly recession dates cannot simply be converted into quarterly recession dates as the monthly and quarterly business cycles are determined separately by the NBER. Rather, the average P_{AES} for a given quarter is used, regardless of the number of Beige Books available in that quarter.

[21] The optimal rule for business cycle dating according to the SPF has not been fully explored, and changing the thresholds or even changing the type of rule used may produce improvements in the discrepancy and lead values.

	NBER Start	NBER Declared	SPF Start	Discrepancy (Quarters)	SPF Declared	Lead (Days)	P_{AES} Start	Discrepancy (Quarters)	P_{AES} Declared	Lead (Days)
Recession	1980-Q2	6/3/1980	1980-Q2	0	Unavailable		1980-Q2	0	5/14/1980	20
	1981-Q4	1/6/1982	1981-Q3	-1	Unavailable		1981-Q4	0	12/16/1981	21
	1990-Q4	4/25/1991	1990-Q4	0	11/28/1990	148	1990-Q4	0	12/5/1990	141
	2001-Q2	11/26/2001	2001-Q4	2	11/20/2001	6	2001-Q1	-1	3/7/2001	264
	2008-Q1	12/1/2008	2008-Q1	0	2/12/2008	293	2008-Q1	0	4/16/2008	229
	2020-Q1	6/8/2020	2020-Q2	1	5/15/2020	24	2020-Q1	1	5/27/2020	12
Expansion	1980-Q4	7/8/1981	1980-Q4	0	Unavailable		1980-Q4	0	12/10/1980	210
	1983-Q1	7/8/1983	1982-Q3	-2	Unavailable		1983-Q1	0	3/23/1983	107
	1991-Q2	12/22/1992	1991-Q3	1	8/21/1991	489	1991-Q2	0	6/19/1991	552
	2002-Q1	7/17/2003	2002-Q2	1	5/21/2002	422	2002-Q1	0	3/6/2002	498
	2009-Q3	9/20/2010	2009-Q3	0	8/14/2009	402	2009-Q4	1	1/13/2010	250
	2020-Q3	7/19/2021	2020-Q3	0	8/14/2020	339	2020-Q3	0	10/21/2020	271

Table 4: Comparison of Quarterly Real-Time Recession Indicators

Notes: Discrepancy shows the number of quarters between the first actual and first perceived quarter of recession or expansion according to the SPF or P_{AES} model model; Lead shows the days between the time at which a recession or expansion is first declared according to the SPF or P_{AES} model model, and the date when the business cycle turning point is officially declared by the NBER; NBER Start dates are the first quarter after the business cycle turning point. The declared dates according to the SPF are the dates on which the SPF was published in the given quarter.

4. Conclusion

There are three points to emphasize in closing. The first is that the sentiment expressed within the Beige Book can be captured using a very simple procedure. Advanced methods or tools for natural language processing are not required to interpret either the concern or optimism communicated by contacts in each report. The lexicon approach to text analysis can be used to create a sentiment index that is reflective of the aggregate economy from the Beige Book text. Additionally, the sentiment indexes created using either general lexicons or economic lexicons capture the broad movements in economic activity well, though they differ in how positive sentiment may be during periods of expansion and how pessimistic contacts seem in the depths of recession as discussed in **Section 2.1**. This is not surprising, though, given the existing literature comparing the performance of generic and domain-adapted lexicons for capturing sentiment.

Second, not only does the Beige Book contain sentiment related to economic developments, but the text also reflects the major current events that influence how contacts feel overall, as shown in **Figure 2**. The basic shock indexes created using word frequency correspond with movements in sentiment, with the largest shocks typically depressing sentiment by a full standard deviation. The inclusion of these shock indexes to account for poor sentiment that is not related to economic developments helps explain some discrepancies between the general and economic sentiment indexes. However, the shock indexes do not fully compensate for the generic nature of the general lexicons used.

Lastly, the combination of the economic sentiment and shock indexes in a probit model produce especially accurate real-time readings of the U.S. business cycle without the addition of other data or economic indicators. The resulting probabilities are comparable to some closely-followed alternative measures, specifically the related question from the Survey of Professional Forecasters and the Sahm rule. And, though the Beige Book may not be as timely as these alternatives, the probabilities obtained using Average Economic Sentiment from the Beige Book would allow an observer to identify business cycle turning points well in advance of the official NBER announcements.

These points do not imply that analysts should comfortably substitute reading the Beige Book with monitoring the Average Economic Sentiment Index. The sentiment scores help quantify the information in the Beige Book to track broad trends within the economy over time, but ultimately the value of the report comes from the nuance of the granular anecdotes that Federal Reserve economists and executives collect from business and community leaders. Neglecting to read the Beige Book and instead electing to focus on only the numerical scores may cause an observer to detect big fluctuations in economic conditions without understanding the driving forces behind those changes. **Section 2.1** highlights where the index misses some of these pivotal details in the latest report and where the lexicon approach to text analysis may more generally mischaracterize a piece of text. Accordingly, the Beige Book sentiment scores and accompanying probabilities should be seen as a tool complementary to the written report and other measures typically used in economic analysis.

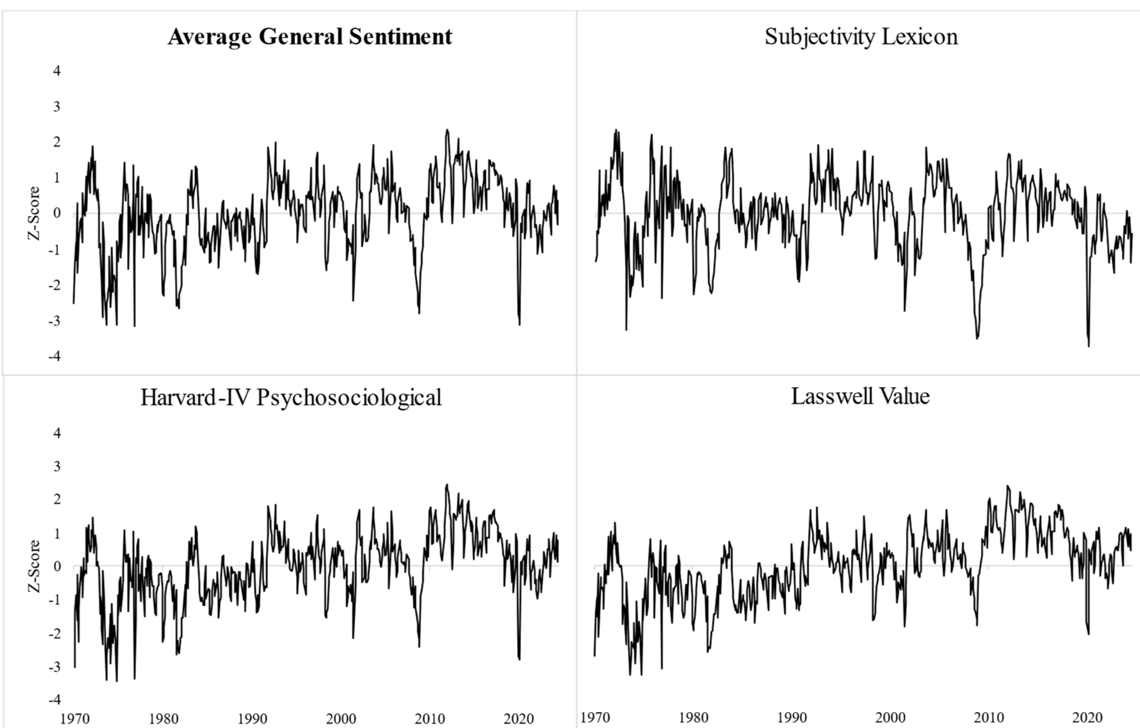
Of course, this tool can be developed further. One possible improvement may be to weigh sector-specific sentiment according to the sector's contribution to gross domestic output or share of employment. Current District Beige Books contain different subsections, some of which relate to specific sectors of the economy (e.g. the energy section of the Dallas report). This has not always been the case, though, and these subsections may change over time within District reports and are not the same across Districts. For these reasons, creating sector-level sentiment may be fairly challenging.

An additional avenue of future research also relates to sentiment analysis more broadly. Confidence intervals exist for the recession probabilities derived from the probit model given a Beige Book sentiment score, but confidence intervals for the sentiment scores themselves do not and this idea is rarely discussed in the existing literature. Obvious potential solutions include calculating sentiment many times according to random subsamples of sentences within the text or according to random subsamples of words within the dictionaries to create a distribution of possible sentiment scores. However, these and other options need to be explored carefully before they are used for drawing inferences.

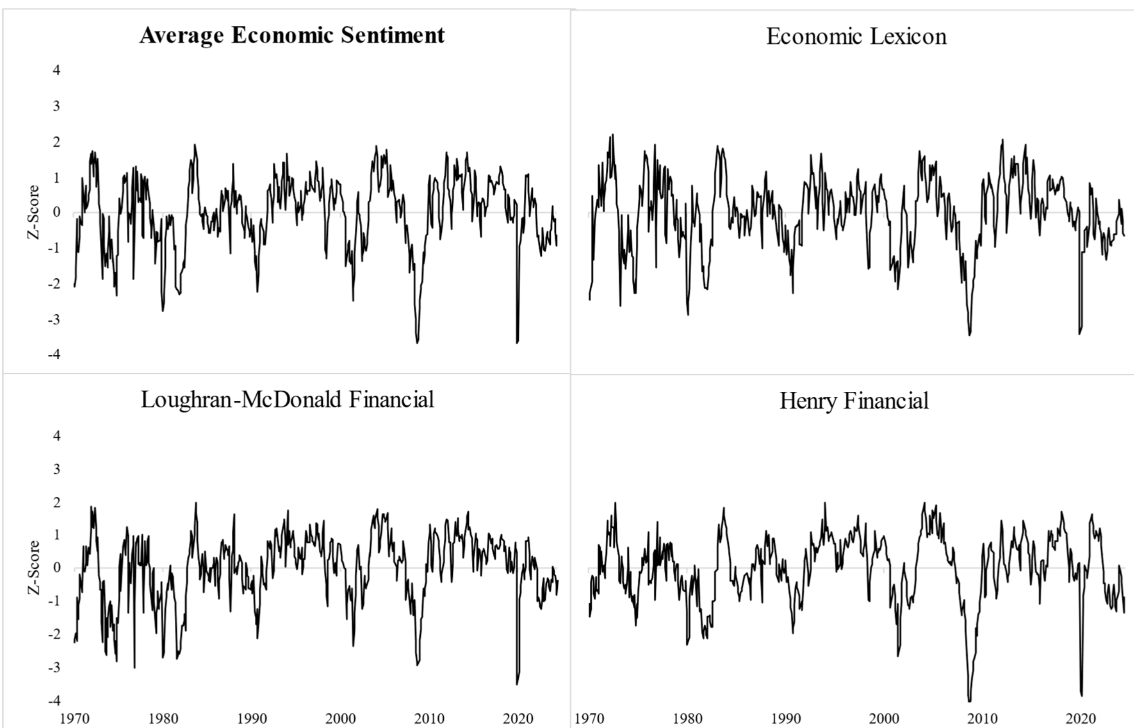
Appendix I: Dictionary Selection

We have focused on using two aggregate sentiment measures to represent sentiment within the Beige Book reports: Average General Sentiment and Average Economic Sentiment. Three lexicons are used in each aggregate measure. The details of the six individual lexicons are grouped together by their category and presented alongside the corresponding sentiment indexes.

1. Average General Sentiment
   a. General Inquirer: Originally developed by Stone et al., (1966) the General Inquirer (GI) contains approximately 9,000 unique word entries. The GI was updated in 2000, and contains positive and negative classifications according to two different sources, the Harvard-IV Psychosociological Dictionary (an updated version of the Harvard-III Dictionary detailed by Stone et al. 1966) and the Lasswell Value Dictionary (described by Namenwirth and Weber 1987). About 6,000 words are shared between the two dictionaries, while the Harvard-IV has an additional 2,000 and the Lasswell dictionary has an additional 1,000. Classifications according to each source are used in creating the Average General Sentiment index and are considered in our analysis to be two distinct dictionaries. Unsurprisingly, the sentiment indexes created using these two individual dictionaries are very similar.
   b. Subjectivity Lexicon: A lexicon of over 8,000 entries, with approximately 93% marked as expressing positive or negative sentiment (Wilson et al. 2005). This is in stark contrast to the GI discussed above, as fewer than half of its unique entries are classified as positive or negative.



2. Average Economic Sentiment
   a. Economic Lexicon: A dictionary of over 5,500 positive and negative words "specifically designed for textual applications in economics." It draws on economic news articles from the U.S. and U.K. published between 1980 and 2020, the monthly Economic Bulletin released by the European Central Bank, and the Beige Books released between 1983 and 2020 (Barbaglia et al. 2022).
   b. Loughran-McDonald Dictionary: Finance-specific sentiment dictionary consisting of over 2,500 positive and negative words developed as an alternative to the Harvard Psychosociological Dictionary, after finding that the general dictionary classifies 73.8% of words as negative that are not considered negative in financial contexts (Loughran and McDonald 2011).
   c. Henry Financial Dictionary: A short list of positive and negative words (fewer than 200) that is shown to be useful in analysis of company earning press releases. The dictionary is used by Henry (2008) to show through a short-window event study that investors are influenced by the tone or sentiment of a corporate earnings press release.



Each aggregate measure is somewhat less noisy than its underlying indexes. More significantly, the economic indexes are less noisy than the general indexes and demonstrate considerable agreement on the depth of negative sentiment during recessions and on the height of positive sentiment during expansions. This is especially interesting considering the wide variety of economic dictionary sizes, ranging from fewer than 200 words to more than 5,000. This may suggest that the exact choice of economic or financial dictionary is not critically important to our lexicon approach to sentiment analysis, as dictionaries of varying robustness produce only marginally different results.

Appendix II: Shock Words

The lists of words have been created to cast a broad net over possible shocks. Arguments can be made against the inclusion of some words in the shock lists (snow, deforestation, erosion, etc.) but broadly speaking the current shock indexes highlight the major extraneous shocks and fulfill their intended purpose. However, the shock indexes will likely be refined over time to minimize noise.

Physical Shock List				
earthquake	storm	extreme temperature	tropical depression	climate change
hurricane	superstorm	mudflow	tsunami	climate
flood	tropical storm	downpour	hail	global warming
tornado	landslide	torrential	snow	deforestation
wildfire	ice storm	extreme weather	wind	polluting
blizzard	cyclone	terrible weather	acid	toxic
fire	heat wave	natural disaster	erosion	erosion
spill	chemical spill	nuclear		

Political Shock List				
impeachment	lawless	war	shooting	police
invasion	assassination	demonstration	election	campaign
protest	strike	crime	terror	tragedy
riot				

REFERENCES

Armesto MT, Hernandez-Murillo R, Owyang MT, and Piger J. 2009. “Measuring the information content of the Beige Book: A mixed data sampling approach.” *Journal of Money, Credit, and Banking* 41: 35–55

Baker SR, Bloom N, and Davis SJ. 2016. “Measuring economic policy uncertainty.” *The Quarterly Journal of Economics* 131: 1593–1636

Balke NS, and Peterson D'A. 2002. "How well does the Beige Book reflect economic activity? Evaluating qualitative information quantitatively." *Journal of Money, Credit, and Banking* 34: 114–136

Balke NS, Fulmer M, and Zhang R. 2017. "Incorporating the Beige Book into a quantitative index of economic activity." *Journal of Forecasting* 36: 497–514

Barbaglia L, Consoli S, Manzan S, Pezzoli LT, and Tosetti E. 2023. "Sentiment analysis of economic text: A lexicon-based approach." *Working paper*. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4106936

Blinder AS. 1997. "Distinguished lecture on economics in government: What central bankers could learn from academics—and vice versa." *Journal of Economic Perspectives* 11: 3–19

Bloom N. 2009. "The impact of uncertainty shocks." *Econometrica* 77: 623–685

Chauvet M, and Piger J. 2008. "A comparison of the real-time performance of business cycle dating methods." *Journal of Business & Economic Statistics* 26: 42–49

Chen Z, Iqbal A, and Lai H. 2011. "Forecasting the probability of US recessions: a probit and dynamic factor modelling approach." *The Canadian Journal of Economics* 44: 651–672

Devlin J, Chang MW, Lee K, and Toutanova K. 2019. "BERT: Pre-training of deep bidrectional transformers for language understanding." *Proceedings of the 2019 Conference of the North American Chapter for the Association for Computational Linguistics: Human Language Technologies.*

Estrella A, and Mishkin FS. 1998. "Predicting US recessions: financial variables as leading indicators." *The Review of Economics and Statistics* 80: 45–61

Federal Reserve Bank of Minneapolis, *Beige Book Archive*: https://www.minneapolisfed.org/region-and-community/regional-economic-indicators/beige-book-archive

Fettig D, Rolnick AJ, and Runkle DE. 1999. "The Federal Reserve's Beige Book: a better mirror than crystal ball." *Federal Reserve Bank of Minneapolis: The Region* 1999: 28–32

Filippou I, Garciga C, Mitchell J, and Nguyne MT. 2024. "Regional economic sentiment: constructing quantitative estimates from the Beige Book and testing their ability to forecast recessions." *Federal Reserve Bank of Cleveland: Economic Commentary* 8: 1–8

Gascon CS, and Martorana J. 2024. "What's behind the recent slump in consumer sentiment?" *Federal Reserve Bank of St. Louis: On The Economy Blog*

Handlan A. 2022. "FedSpeak matters: Statement similarity and monetary policy expectations." *Working Paper*. Available at https://handlanamy.github.io/MyFiles/Handlan_fedspeak.pdf

Handlan A. 2022. "Text shocks and monetary surprises: Text analysis of FOMC statements with machine learning." *Working Paper*. Available at https://handlanamy.github.io/MyFiles/Handlan_TextShocks_sub.pdf

Gati L, and Handlan A. 2023. "Monetary communication rules." *Working Paper*. Available at https://handlanamy.github.io/MyFiles/GatiHandlan_MonetaryCommunicationRules_2023.pdf

Hao L, and Ng ECY. 2011. "Predicting Canadian recessions using dynamic probit modelling approaches." *The Canadian Journal of Economics* 44: 1297–1330

Huang AH, Wang H, and Yang Y. 2023. "FinBERT: A large language model for extracting information from financial text." *Contemporary Accounting Research* 40: 806–841

Henry E. 2008. "Are investors influenced by how earnings press releases are written?" *Journal of Business Communication* 45: 363–407

Loughran T, and McDonald B. 2011. "When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks." *Journal of Finance* 66: 35–65

Namenwirth JZ, and Weber RP. 1987. *Dynamics of culture*. Boston, Massachusetts. Published by Allen and Unwin. Accessed through https://inquirer.sites.fas.harvard.edu/lasswell.htm

Proellochs N, and Feurriegal S. 2016. "SentimentAnalysis." R package version 1.3-5. https://cran.r-project.org/package=SentimentAnalysis

Rogers A, Kovaleva O, and Rumshisky A. 2020. "A primer in BERTology: What we know about how BERT works." *Transactions of the Association for Computational Linguistics* 8: 842–866

Sahm C. 2019. "Direct stimulus payments to individuals." Boushey H, Nunn R, and Shambaugh J. *Recession Ready: Fiscal Policies to Stabilize the American Economy* 67–92, The Brookings Institution

Stone PJ, Dunphy DC, Smith MS, and Ogilvie DM. 1966. *The general inquirer: A computer approach to content analysis*. Cambridge, Massachusetts, and London, England. Published by The M.I.T. Press

Wilson T, Wiebe J, and Hoffman P. 2005. "Recognizing contextual polarity in phrase-level sentiment anlysis." *Proceedings of Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing* 347–354

Wright J. 2006. "The yield curve and predicting recessions." *Board of Governors of the Federal Reserve System: Finance and Economics Discussion Series 2006–2007*